Naval Research Laboratory

Washington, DC 20375-5320

NRL/FR/5580--96-9826

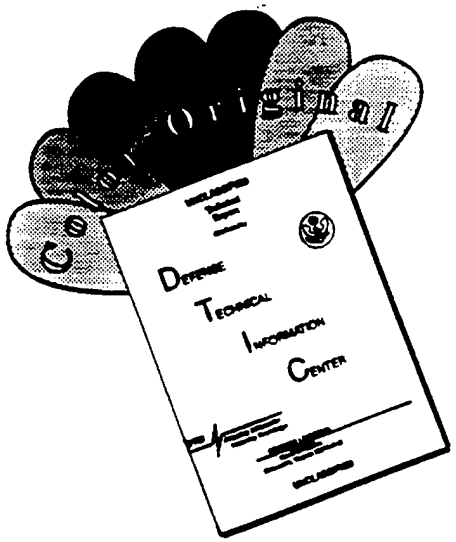
Underwater Imaging with Acoustic Lenses: Image Processing and Visualization

BEHZAD KAMGAR-PARSI

Advanced Information Technology Branch
Information Technology Division

August 30, 1996

DISCLAIMER NOTICE



THIS DOCUMENT IS BEST QUALITY AVAILABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF COLOR PAGES WHICH DO NOT REPRODUCE LEGIBLY ON BLACK AND WHITE MICROFICHE.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave Blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	August 30, 1996	

4. TITLE AND SUBTITLE

Underwater Imaging with Acoustic Lenses: Image Processing and Visualization

5. FUNDING NUMBERS

PE - 62315N

6. AUTHOR(S)

Behzad Kamgar-Parsi

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Naval Research Laboratory
Washington, DC 20375-5320

8. PERFORMING ORGANIZATION REPORT NUMBER

NRL/FR/5580--96-9826

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Office of Naval Research
Arlington, VA 22217

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

This report discusses image processing, scene reconstruction, and visualization techniques used for underwater acoustic images taken with lens-based systems from a stationary platform. These systems are designed for high-resolution imaging of objects from distances of a few meters. The acoustic lenses used for beamforming in the imaging systems are made of crystal polystyrene and are cut in cylindrical or spherical shapes. The cylindrical lens has a fan-shaped beam pattern and produces a 2D intensity image or shadowgram, while the spherical lens with a cone-shaped beam produces a 3D intensity image. Test images obtained by the spherical lens contain a remarkable degree of detail.

14. SUBJECT TERMS

Acoustic imaging
Acoustic lens
Image processing
Underwater imaging
Visualization

15. NUMBER OF PAGES

30

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std 239-18
298-102

CONTENTS

UNDERWATER IMAGING WITH ACOUSTIC LENSES: IMAGE PROCESSING AND VISUALIZATION

INTRODUCTION

In a joint effort sponsored by the Office of Naval Research, the Naval Explosive Ordnance Disposal Technology Division, the University of Washington Applied Physics Laboratory, the ARINC Research Corp., and the Naval Research Laboratory have developed prototype systems for high-resolution, closeup, underwater acoustic imaging. This report discusses the image processing, scene reconstruction, and visualization techniques developed, implemented, and evaluated during the course of this project and also presents a sample of the acoustic images.

The two imaging systems discussed in this report are based on an acoustic lens made of crystal polystyrene. One of the imaging systems has a *line focus* with fan-beam patterns and generates 2D intensity images, or shadowgrams. The other has a *point focus* with conical-beam patterns and produces 3D range and intensity data. Hence, the two imaging systems require different processing and visualization approaches. For 2D intensity images, a large number of well-understood processing techniques exist in the literature, and displaying them is straightforward. For 3D range/intensity images, however, many problems in target reconstruction and visualization, and even in filtering, are not yet resolved in a satisfactory manner and are current topics of research.

The imaging system prototypes used in this work have a single transducer in the focal plane, and image acquisition is done by mechanical scanning. Thus, the images represent an ideal that is free of acoustical and electrical cross-talk noise. In the next generation imaging systems currently under development, focal planes will be populated with a dense array of transducers. Therefore, cross-talk noise will be present in the images, and their reconstruction and processing will be more challenging. A motivation for considering these two systems is to investigate the tradeoff between design simplicity and image quality: the conical-beam system is more difficult to build and requires more image processing; on the other hand, it produces sharper, more informative images.

The imaging systems are meant to be diver-held. Using an acoustic lens eliminates the extensive electronics of traditional beam-forming systems, thus the imaging system will be more compact and consume less power. Moreover, the processing has to be done on-line. Therefore, we have emphasized processing and visualization techniques that are reasonably fast, rather than those techniques that would perform best. Furthermore, we have concentrated on processing algorithms that are useful for a general scene rather than model-based algorithms that would perform well in specific scenes.

This report first discusses the fan-beam imaging system, then the conical-beam imaging system, and concludes with remarks on areas that may need further investigation.

Manuscript approved May 28, 1996.

FAN-BEAM-LENS

The Imaging System

In this imaging configuration, beamforming is accomplished with a solid cylindrical lens made of crystal polystyrene. A curved transducer element, 2 mm tall by 6 cm long, is placed at the on-axis focal point behind the lens and is resonant at 3 MHz. The lens has an anechoic rubber coating with a 0.57 cm slit to block azimuthal sidelobes. Together the element and the lens form a fan beam 0.17° in elevation and 4° in azimuth at the -3 dB points at the target distance. To obtain an image, the system mechanically scans and interrogates the object at 0.35 cm intervals top to bottom and 7.8 cm intervals left to right. (Scanning at 7.8 cm intervals was done because the 4° beam was not wide enough to ensonify targets of interest in their entirety.) This scanning process typically takes 15 min. but simulates a lens with good off-axis performance coupled with a 150 × 7 element array with no interelement cross-talk.

Image Processing and Display

Displaying the data taken by the fan-beam imaging system is straightforward. The data are in a grid of cross-range and range (or time delay). We form a regular rectilinear grid (or a polar grid, depending on the scanning strategy used), appropriately scaled so that range and cross-range have the proper dimensions. We also digitize the acoustic backscatter energies to the usual 8-bit gray level range, i.e., 256 discrete values, and assign a color palette to the intensities to highlight contrast. Thus, we obtain a 2D image where the intensity is the backscatter energy.

For the data sets in which the range and cross-range are scanned with different resolutions, e.g., range is sampled at 1 cm intervals while cross-range is sampled at 0.5 cm intervals, the grid size is selected to be the smaller sampling interval. We use a bilinear interpolation to fill in the grid sites from neighboring points. Also, to obtain better contrast for the middle intensities, we transform the backscatter intensities to the 8-bit gray level range such that the n-th lowest and the n-th highest backscatter intensities (with $n \approx 10$) coincide with 0 and 255, respectively, rather than setting the lowest and the highest equal to 0 and 255.

These images, or shadowgrams, typically consist of a few bright spots in a mostly dim and noisy background. The bright spots come from surface patches that have a surface normal pointing directly toward the imaging system lens. These include such features as corner reflectors, bolt heads, rivets, and cylindrical pipes. Much of the information about the target of interest, however, is contained in the dim background. The shadowgrams can be enhanced through fairly standard and straightforward image processing techniques. To brighten the background, we perform a nonlinear scaling of the image intensities. We do this by expanding the range of lower intensities while compressing the range of high intensities. In this way, the dim background becomes brighter and acquires greater contrast. Typically, we move the average image intensity to the middle of the gray level range, i.e., to 127. The solid line in Fig. 1 shows a typical nonlinear transform of the backscatter intensities to the 0 to 255 gray level image.

The effect of background noise is reduced by low pass filtering or local weighted averaging (e.g., convolving the image with a Gaussian kernel). A similar, and somewhat better, effect is achieved through median filtering, which does not blur object edges while smoothing the high frequency noise. Median filtering, however, is computationally more costly than averaging.

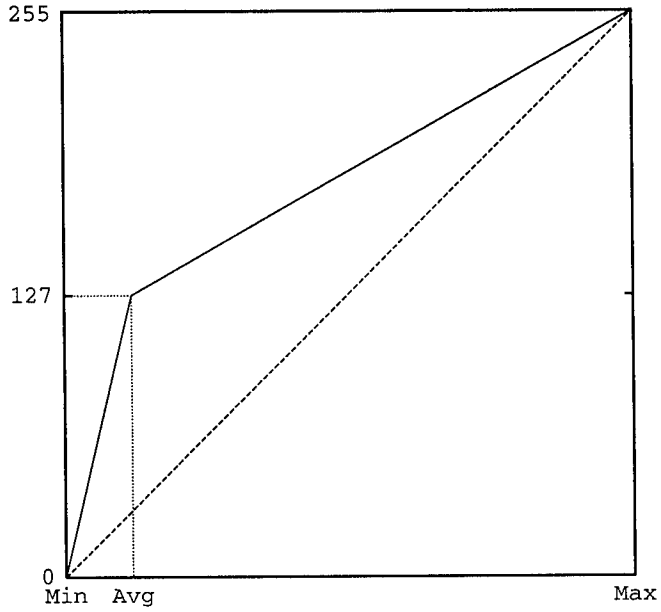


Fig. 1 — A nonlinear transform of backscatter energies to the 8-bit gray-level, enhancing weak-return contrasts

We also implemented and evaluated a variety of standard techniques for enhancing 2D images. None were as useful as the techniques described above. One such technique, which is commonly applied to optical images, is computing the gradient image, or adding it (with various weights) to the original image [1,2]. This technique is intended to enhance perception of object edges. However, given the additional processing time and the modest perceptual improvements, we concluded that this filter as well as many others was not useful for the shadowgrams.

Experiments

The fan-beam imaging system mimics a vertical array of closely packed transducers with a beam that is narrow in elevation and wide in azimuth. The fan-beam data are the backscatter energy as a function of range and elevation, which can be naturally represented as a 2D gray scale image. To test the imaging system, four different objects (Figs. 2 through 5) were placed in a water tank.

Object A (Fig. 2) is a circular plate (45 cm in diameter) with four circular substructures (ranging between 10 cm and 18 cm in diameter) as well as four ribs (0.6 cm wide by 8.5 cm high) perpendicular to the plate.

Object B (Fig. 3) is the head of a cylinder (47 cm in diameter) that has a circular groove (18 cm in diameter) enclosing eight nuts (1.9 cm in diameter) and a raised circular area (8.2 cm in diameter and 0.5 cm high). The other half of the drum head is a flat surface that makes a 25° slope with the first half plane containing the bolts.

Object C (Fig. 4) is a section of a cylinder (47 cm in diameter) with a recessed well (18 cm in diameter and 10 cm deep) that has eight hexagonal nuts (1.9 cm in diameter) and bolts.

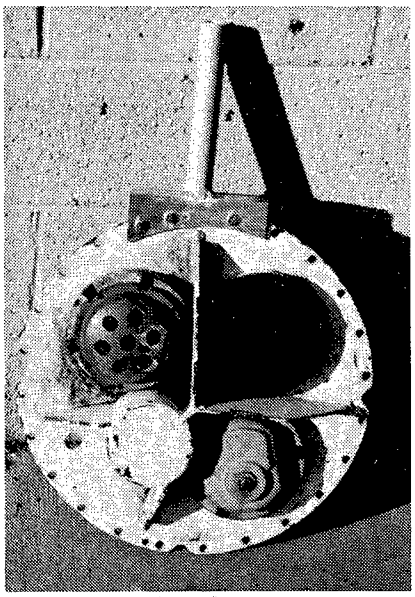


Fig. 2 — Photograph of object A, a plate with four circular substructures

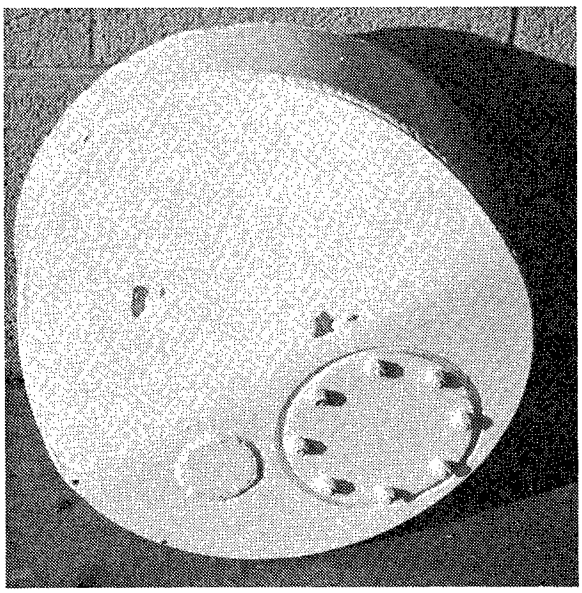


Fig. 3 — Photograph of object B, a drum head with a sloped surface

Object D (Fig. 5) is a 14-in.-long cylinder with a diameter of 9 in. with one end open and the other end closed. The cylinder has two sets of six hexagonal nuts (diameter of 1/2 in. and height of 3/8 in.) and bolts, and a 2-in.-tall fin.

We present a sample of the images taken under two imaging scenarios: objects suspended in a water tank and objects resting on a sandy bottom. Very little processing has been performed on the fan-beam images shown here. They consist of rescaling the intensities so that the average intensity is 127, which is midway between zero and the maximum intensity of 255 (see Fig. 1). This helps to bring out features that have low backscatter energy, while at the same time enhancing the effects of noise. We have also used median filtering (with a size 7 window) to reduce the effect of noise. The

Fig. 4 — Photograph of object C, a cylinder with a recessed well

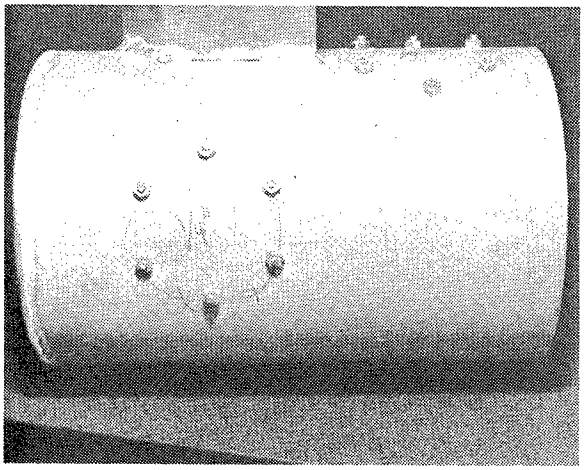


Fig. 5 —Photograph of object D, a foot-long cylinder with two hexagonal bolt patterns and a fin

advantage of median filtering over local averaging is that median filtering tends to preserve intensity edges in the image, although it is more time consuming. Images have been enlarged using bilinear interpolation so that the long axis is 256 pixels. The typical CPU time for processing and displaying these images is about 20 s on a SPARC 10/30 workstation. In the images presented in Figs. 6 through 9, the horizontal axis is the range, the vertical axis is the elevation, and the backscatter energy is shown by the image intensity. Color is used to highlight the intensity: red is high, green is low.

Suspended Objects

In this imaging scenario, the objects are suspended in a water tank away from the tank bottom and walls, which are strong backscatterers.

Figure 6 shows the fan-beam image of object A at three different angles of incidence. The angles between the ensonification direction and the normal to the object plate are, respectively, 30°, 45°, and 60° from left to right. These images show considerable detail, not only the general shape of the object. Note, for example, the bolt heads used to attach the plate to the bar, and the rivets in the upper

Fig. 6 — Fan-beam images of object A at three different angles of incidence: 30°, 45°, and 60°. The object is suspended in the water tank.

right quadrant of the plate. Also note that in the 30° image, the details of the structure in the lower right quadrant are quite visible, while in the 60° image, the hemispherical feature in the upper left quadrant, including the holes, can be easily discerned. The vertical high-intensity line present in the bottom of all three images is caused by the strong acoustic return from the corner formed by the rib and the plate. Also note the reverberations of the hollow bar attached to the top of the plate, which appear as a series of high-intensity lines parallel to the tube and spaced at regular range intervals.

Figure 7 shows the fan-beam image of object B at a 45° angle of incidence. The front rim of the cylinder and the circular bolt pattern are quite visible. The small circular relief and the groove around the bolt pattern are also noticeable, while the sloped plane is not seen at all because the angle of incidence is so large that the acoustic returns are insignificant. The high intensity lines at the top of the image are due to the return (and its reverberations) from the bar holding the drum head.

Figure 8 shows the image of object C with the recessed well facing the lens. The plate of the recessed well, the bolt heads, and the mouth of the well are clearly visible. The bright spot and trail in the middle of the image is due to the return (and its reverberations) from a hollow bar that is holding the cylinder in the tank. The bar is perpendicular to the image plane.

Figure 6 was obtained from a distance of 1.2 m, while Figs. 7 and 8 were taken from a distance of 1.9 m. For an object at a 2-m distance, the scanning intervals used result in oversampling by about a factor of 2 relative to sampling at the -3 dB down points.

Objects on the Bottom

In this scenario, objects are placed on a layer of sand on the tank bottom. The sandy bottom is a strong acoustic reflector. The lens is placed 30 in. above the floor, about 60 in. from the objects, and looking downward at an angle of about 20°.

Fig. 7 — Fan-beam image of object B at a 45° angle of incidence. The object is suspended in the water tank.

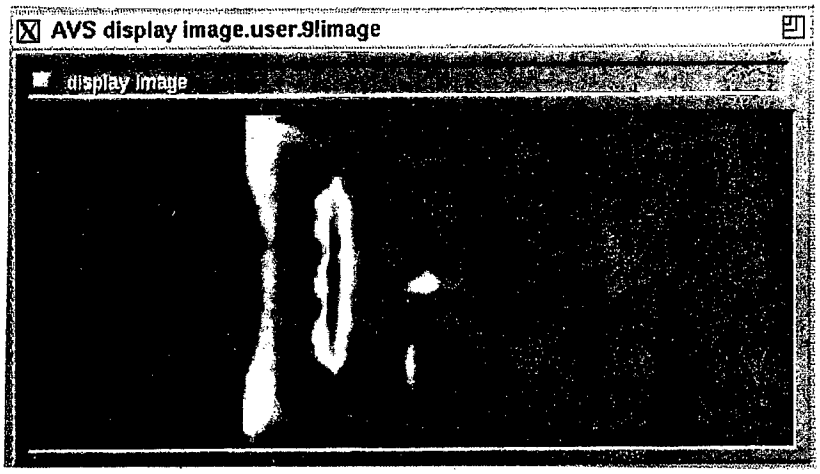


Fig. 8 — Fan-beam image of object C with the recessed well facing the lens. The object is suspended in the water tank.

Figure 9 shows images of objects A and D resting on the sandy bottom. The top-left panel is object D. The other panels are object A at different poses. The strong returns from the bottom and the interreflections between the bottom and the object not only obscure the details of the objects; they even obscure the targets and make them unrecognizable. The mirror effect produces a number of spurious ghost surfaces that appear at longer distances from the lens (which is on the left-hand side of the images), because the returns due to interreflections arrive later than direct reflections. For example in the top-left panel, the bright vertical lines clearly show the direct reflection from the cylinder (the linear surface patch directly facing the lens) and several interreflections between the bottom and the cylinder. Those parts of the cylinder whose surface normal makes an angle with the lens axis reflect little energy back to the lens. Images of object A, which is more complex than the cylinder, do not show a clear mirror effect; rather, object A's features image as a noisy, bright region. The mirror effect makes image interpretation considerably more difficult.

Discussion

Fan-beam images of objects suspended in water have sufficient contrast that a viewer can recognize the object of interest. However, when the object is resting on the bottom, or is close to another reflecting surface that has a reflectivity comparable to that of the object of interest, then the contrast may not be sufficient to discern the object. This is because the backscatter is averaged over a relatively wide angle. Furthermore, interreflections make image interpretation more difficult.

In this phase of the project, we have been concerned with image processing techniques for the general enhancement of images. There exist model-based (or shape-based) filters that are tailored to finding and enhancing given patterns in noisy images with missing parts (shadows). These morphological operations, or template matching techniques, are designed to find and enhance a user-specified shape, e.g., circles or hexagons of a given size, or a given pattern of points. Template matching and morphological operations may be added in the next phase to aid in finding certain features that are important in identifying the type of mine.

Humans are adept at inferring 3D shape of objects from shading information in 2D images. Motivated by this fact, techniques have been developed in computer vision (the so-called shape-from-shading techniques) to reconstruct the 3D scene from shading information in a 2D image. These techniques work under restrictive assumptions, such as Lambertian surface and constant albedo, and attempt to find the illumination direction and surface normals at the same time. They involve solving a set of partial differential equations and are computationally intensive [3]. We are developing shape-from-shading techniques from acoustic shadowgrams, which take advantage of the knowledge of the illumination (ensonification) direction. Knowing the illumination direction makes the solution more robust and reduces the computational cost. The drawback of working with shadowgrams is that shadows (except for their boundaries) do not contain any shape information, while shading in optical images provides shape information in the form of surface normals.

CONICAL-BEAM LENS

The Imaging System

Here we give a brief description of the imaging system; References 4 and 5 provide greater detail. In this imaging sonar, beamforming is done with a bi-concave solid lens made of crystal polystyrene. This lens is lighter than the liquid-filled lens [6] and maintains focus over a wider temperature range. The drawback is the off-axis aberration, which is nonexistent in the liquid-filled lens. At the on-axis focal point behind the lens is a single 3 MHz transducer element 3 mm in diameter. The element and

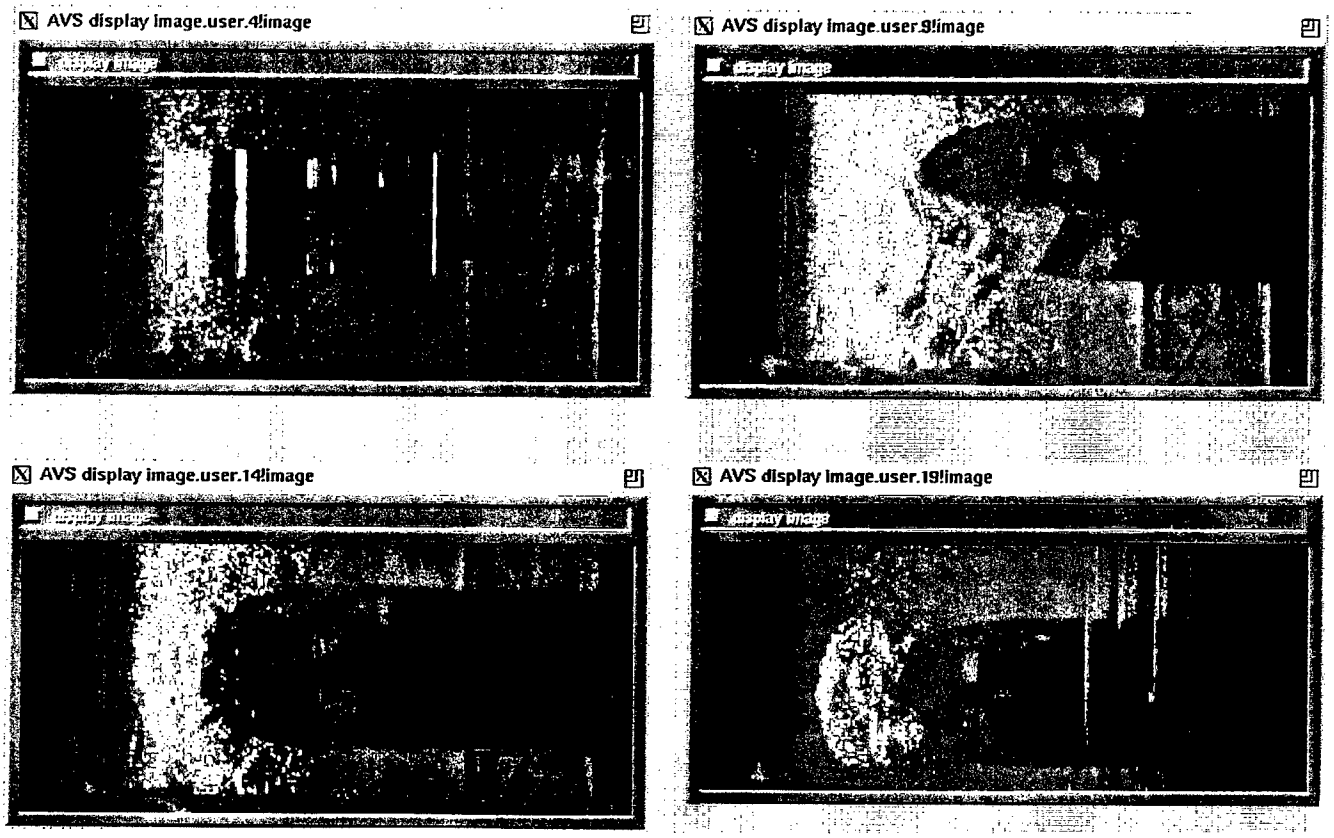


Fig. 9 — Fan-beam images of objects resting on the sandy bottom of the water tank. Top-left panel is the image of object D, while the other three panels are images of object A at three different poses.

the lens together form a conical beam with 0.15° resolution between the -3 dB points at the target distance. To obtain an image, the system mechanically scans and interrogates the object at regular intervals (typically 0.5 cm) from left to right and from top to bottom. This scanning strategy requires a few hours to produce an image, but it mimics an ideal lens with a large focal plane populated with some 100 × 100 transducers, obviously with no aberration and no cross-talk among the elements. To speed up image acquisition, mechanical scanning must be replaced by a focal plane populated with many transducers. However, this will result in significant off-axis image distortions. Techniques are currently under development for reducing aberrations through compound lens design, much as in sophisticated cameras. In addition, innovative scanning strategies are being developed.

Object Reconstruction and Visualization

Data acquired by the sensors consist of the acoustic backscatter energy from within volume elements, or range bins, shaped like sliced cones. The range bins vary in size and may be irregularly positioned in space. To reconstruct the scene, we represent the space with a regular grid of small cubic volume elements, or voxels, and estimate the backscatter from each voxel. A high voxel value indicates the presence of a surface. This is done in two stages. First, since each range bin intersects several voxels, its backscatter must be distributed among them. Second, a voxel may be ensonified by several beams, resulting in different estimates that must be combined to yield a single value. Algorithms for distributing returns involve modeling the beam patterns and computing range bin intersections with voxels. This is conceptually straightforward but can be computationally intensive.

Algorithms for combining estimates are nontrivial because we have to deal with occlusions, partial illuminations, and differences in angle of incidence, particularly if the sensor, or the target, is moving.

Having obtained a 3D gray level image, we must find a way to see the object of interest through the fog of ambient noise. How best to enhance and visualize the object is also a topic of ongoing investigation. There are two main approaches to visualizing 3D scalar data: volume rendering and surface rendering [7]. In volume rendering, we assign to each voxel a color and a partial transparency, and then form images by blending together colored, semitransparent voxels that project to the same pixel on the image plane. In surface rendering, we first apply a surface detector to the volume data, then tile the surface with polygons, and finally render the surface. The main drawback in surface rendering is that we have to classify each voxel as belonging to a surface or not. Its advantage is in rendering speed, particularly if we want to rotate the computer image of the scene and view it from several different angles. This is because the surface voxels are determined and tiled with polygons only once. Rotating and rendering the reconstructed surfaces for subsequent views require little additional CPU time. In volume rendering, however, when the view direction is changed, all the computations must be repeated anew. Nevertheless, more efficient and clever ways of volume rendering are being developed and tested [8].

In our experience, when data are high resolution and relatively clean, it is best to use surface rendering to view the target. Therefore, we first apply a surface detection algorithm to identify the surface voxels; then tile the surface with voxel-sized polygons, which performs a subvoxel interpolation; and finally render the surface. The images shown in this section are produced with this approach. For surface detection, we used the technique described in Ref. 9. (Although, taking the location of the first strong return along each beam or the first maximum return as a surface voxel produced similar images with comparable quality.) For tiling the detected surface with polygons, we used the marching cubes algorithm [10]. For rendering the polygon-tiled surfaces, we used Phong shading [11], which gives the objects a realistic appearance [12].

Our previous work used a 300 kHz liquid-filled lens, which has a lower resolution than the lens used in this project. There, to obtain more detailed, denser images, the lens was mounted on a slowly moving platform. The scene was reconstructed by combining many snapshots taken during a pass over the object of interest (as well as by combining multiple passes over the scene). In this manner, we were able to obtain higher resolution scene reconstructions. Nevertheless, because the data were blurred by sensor motion as well as slight registration errors, we found that surface detection often failed to produce the best possible image. Rather, volume rendering produced 3D images that appeared to be more similar to the objects at hand [6,13].

Experiments

The conical-beam imaging system mimics a 100×100 square array of closely packed transducers with narrow beams. The data is in the form of a 3D grid of backscatter intensity values. To reconstruct and visualize the object of interest, we perform the following operations: grid the data into voxels; apply a surface detection algorithm to label voxels as surface or not; construct a surface by tiling it with small polygons; and, finally, render the surface to obtain a 3D view of the object. To test the imaging system, five different objects, (A, B, C, D (Figs. 2 through 5), and E (Fig. 10)) were placed in a water tank. Object E is a rectilinear grid of bars with variable spacing, which is used to test the resolving power of the imaging system.

The following samples of the images were taken under two imaging scenarios: objects suspended in a water tank, and objects resting on a sandy bottom. The 3D scene is typically a rectilinear grid of $100 \times 100 \times 50$ cubic voxels. The scene reconstruction and visualization of these 3D data takes about 2 min. of CPU time on a Stardent (Kubota) computer.

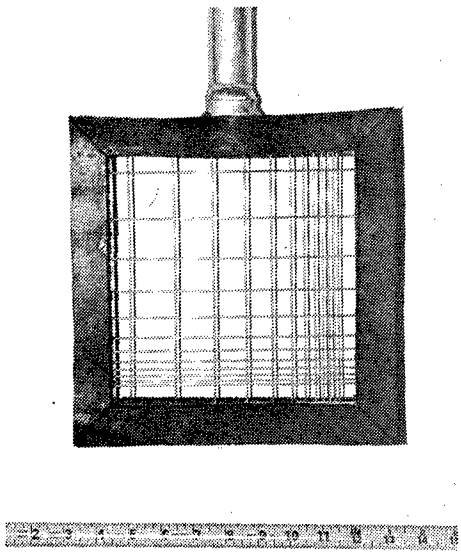


Fig. 10 — Photograph of object E, a 12 × 12 grid with decreasing spacing

Suspended Objects

In these experiments, objects are suspended in a water tank away from the tank bottom, the walls, and other strong reflectors. For suspended targets, the backscatter from the object is much stronger than the ambient background noise, hence surface detection algorithms perform well. These images require little processing to enhance them. Although we used a surface detection algorithm for object reconstruction, similar results were obtained (at a reduced computational cost) by marking the location of the first strong return, or the maximum return, along each beam to be a surface voxel. The objects are imaged from a distance of about 2 m.

Figure 11 shows two surface-rendered views of the grid (object E) shown in Fig. 10. The grid is 12 × 12; the acoustic lens is able to resolve the largest 8 × 8. This corresponds roughly to a resolution of 0.5 cm, which agrees with the expected theoretical resolution.

Figure 12 is the surface-rendered acoustic image of object B. The image was taken with the plane of the bolt pattern facing the lens. The isolated spot to the right is a small portion of the bar used to hold the object in the tank. Note that the bolt pattern, the circular groove, and the small circular relief are clearly visible. The noisy appearance of the sloped surface is partly because of aliasing due to scene quantization and partly because of the unfavorable look angle resulting in weak backscatter and, therefore, less certain surface detection. Aliasing is an artifact of surface rendering of digitized surfaces [14]. In volume rendering, aliasing effects are less pronounced [8].

Figure 13 is an image of object C. The object was posed such that the bottom of the recessed hole directly faced the lens. Note that not only all the bolt/nut combinations clearly stand out, but that even certain bolts and nuts can be distinguished. The noisy appearance of the cylindrical part is again due to unfavorable angle of incidence, weak returns, and aliasing.

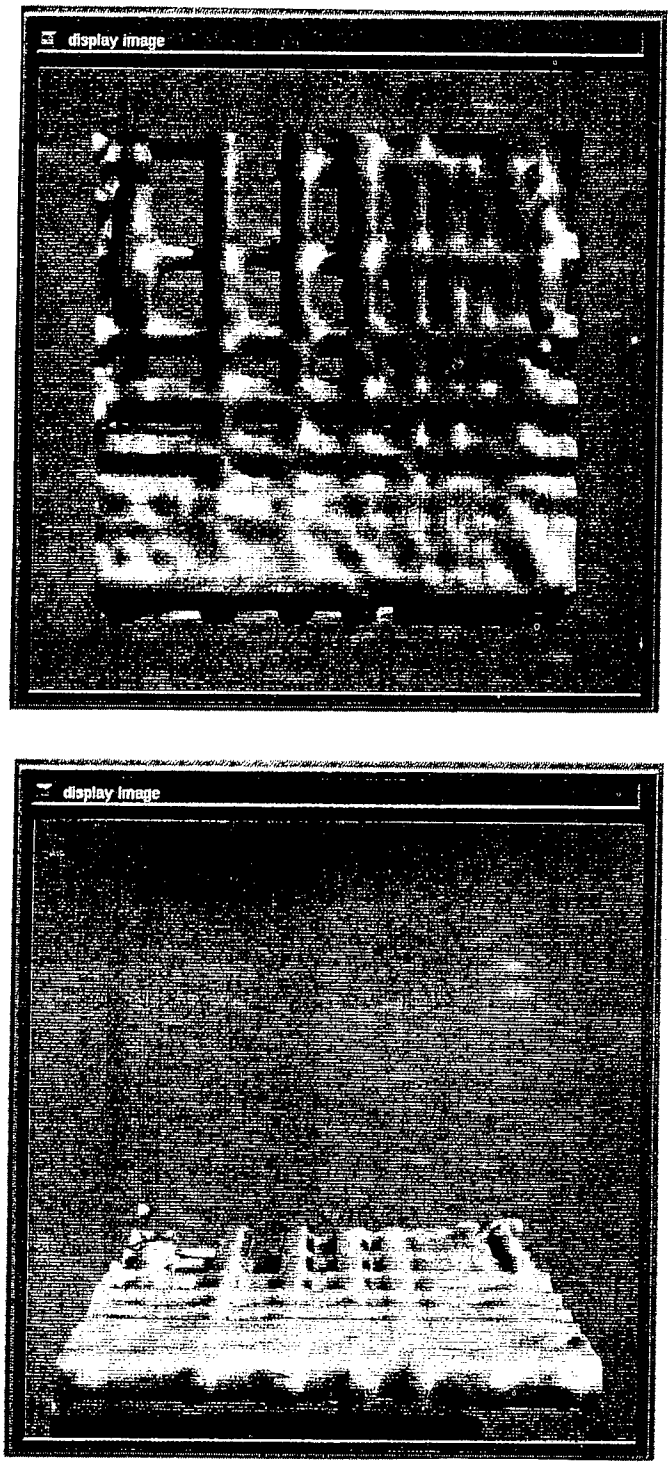


Fig. 11 — Two views of the reconstruction of object E from data taken by the conical-beam lens

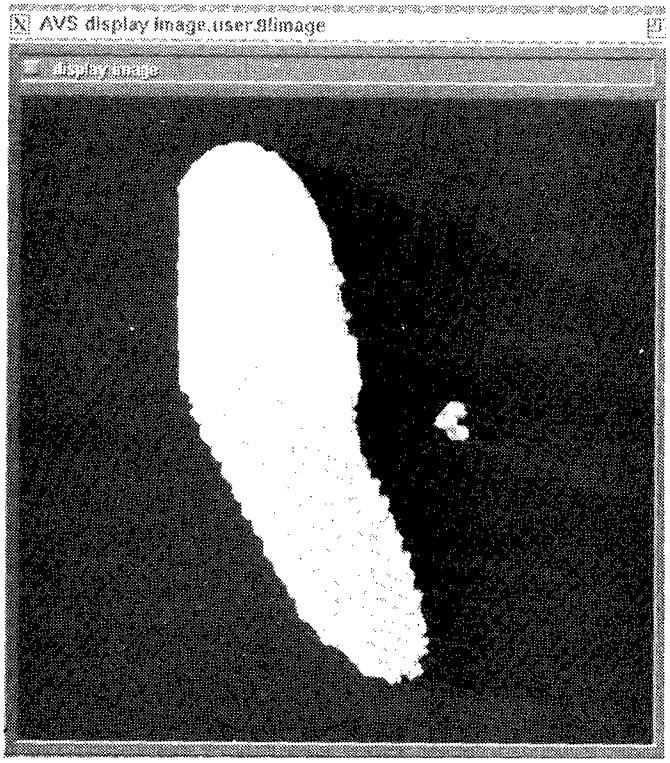


Fig. 12 — A 3D surface rendering of object B from conical-beam lens data

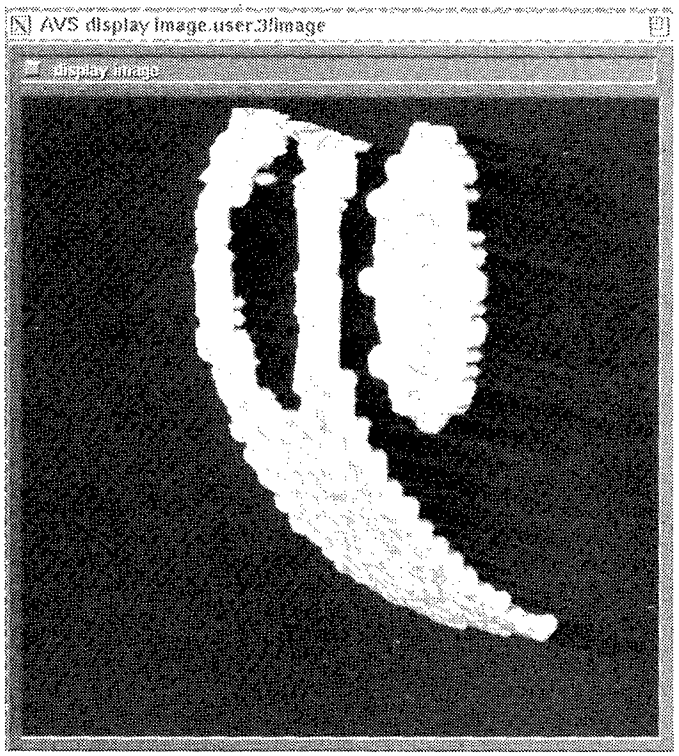


Fig. 13 — A 3D surface rendering of object C from conical-beam lens data

Figure 14 is an acoustic image of object A. The object is directly facing the lens. Note the remarkable degree of detail that can be observed in this image, such as the small hole in the bottom-left quadrant, the rivets, and the three bolt heads on the handle holding the plate. Figure 15 is also an image of object A, this time facing the lens at a 30° angle. Because the plate makes an angle with the lens axis, the angle of incidence is not favorable and the returns are weaker. This results in a less accurate surface detection and, hence, a noisy appearance in which some of the details become difficult to discern. However, one has no difficulty identifying an image as object A. The noisy appearance of the object in Fig. 15 can be improved by applying a set of filters to the 3D data before surface rendering. The filtering process is described in the Appendix. The result of filtering the object rendered in Fig. 15 is shown in Fig. 16. It can be seen that the noisy appearance of the surface is much improved. While in Fig. 16(a) (created from an isosurface with a low value) the larger features are seen clearly, in Fig. 16(b) (created from an isosurface with a higher value) even certain details such as rivets and bolt heads can be discerned.

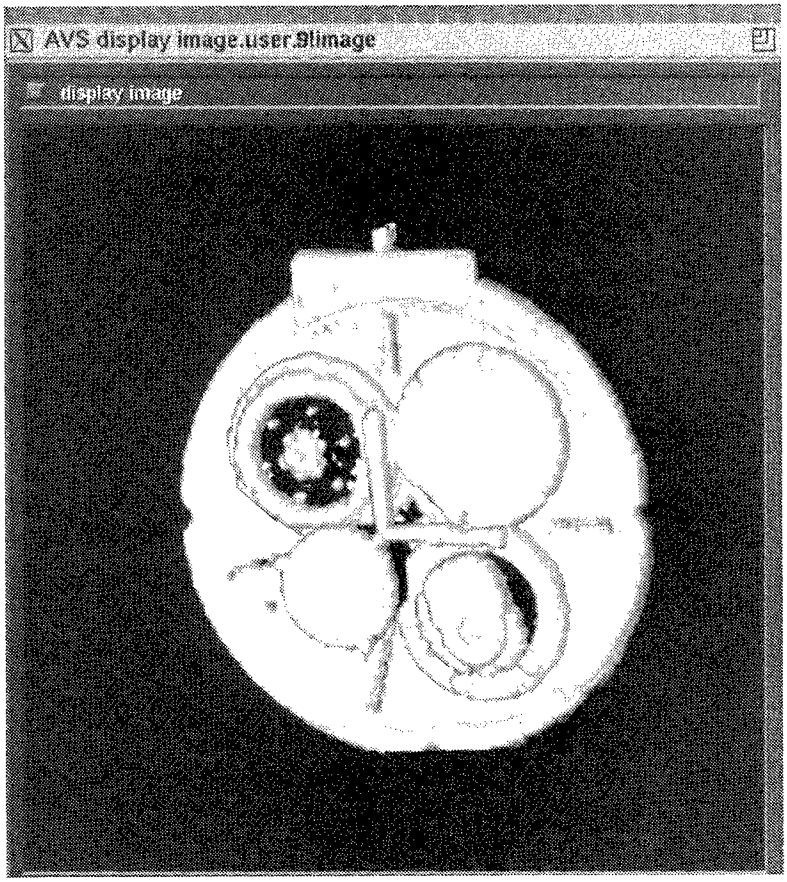


Fig. 14 — A 3D surface rendering of object A from conical-beam lens data, taken in a pose where the plate faces the lens

Objects on the Bottom

In this imaging scenario, objects are resting on a layer of sand on the tank bottom. The sandy bottom is a strong acoustic reflector. The lens is about 30 in. above the floor and 60 in. from the objects, and is looking downward at an angle of about 20°.

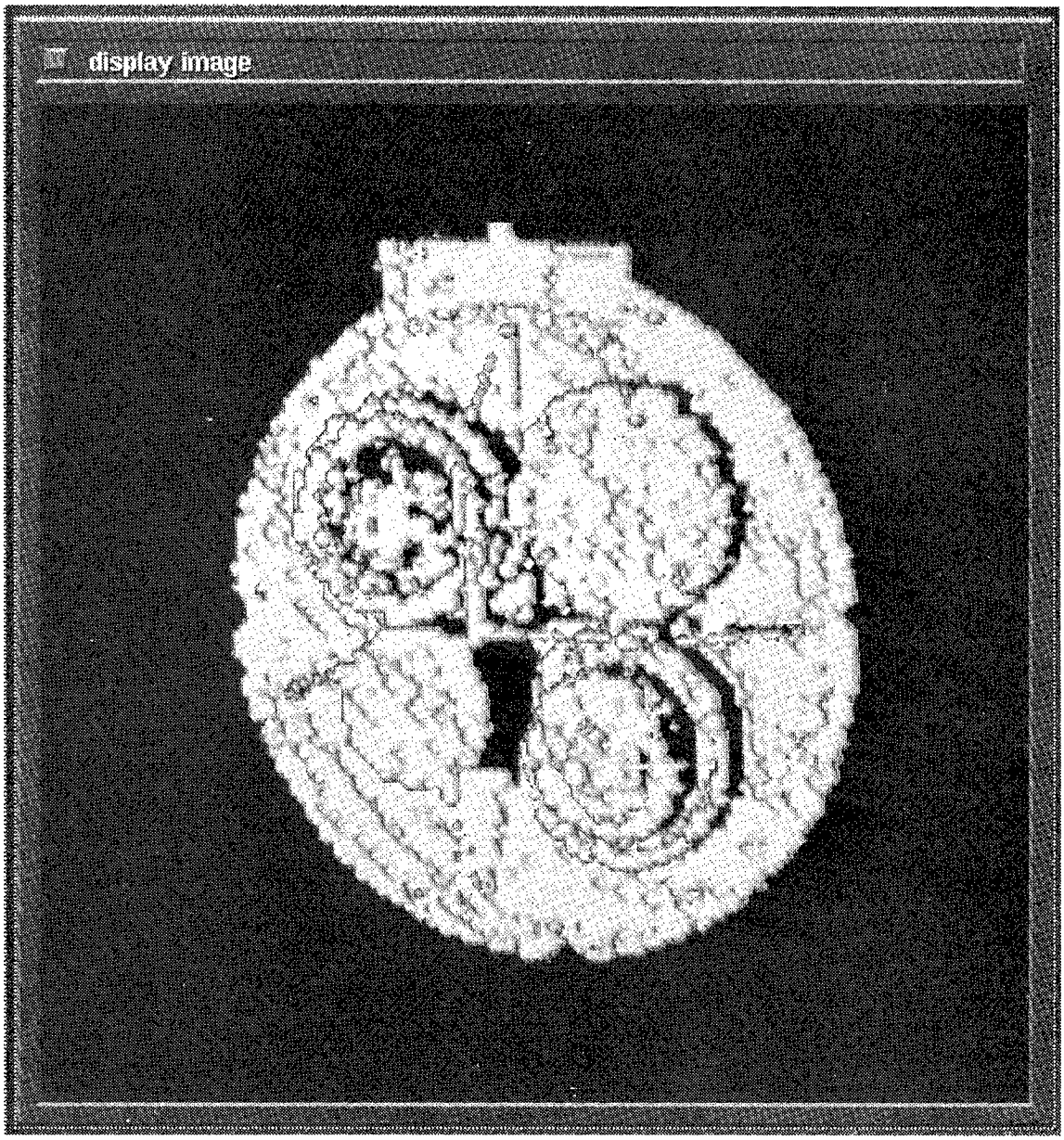
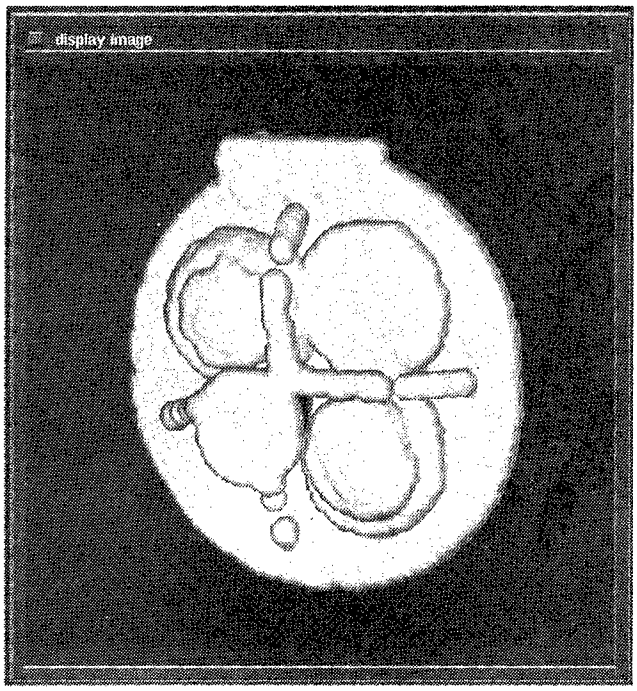
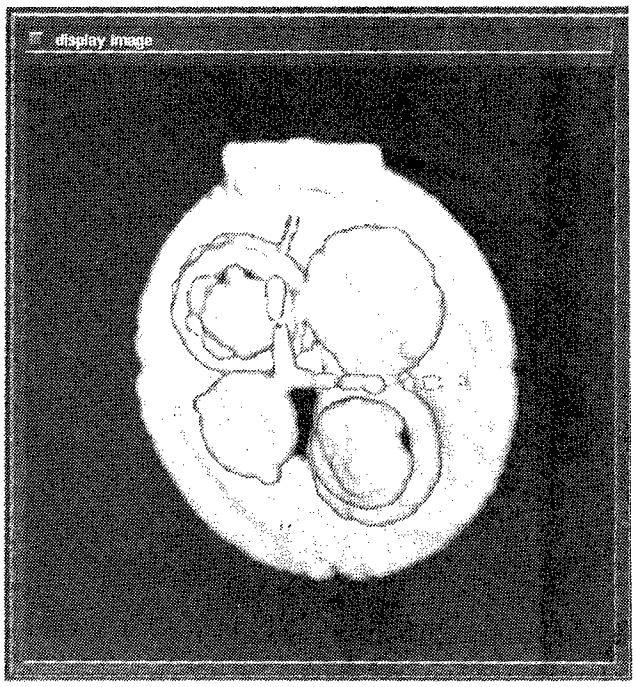


Fig. 15 — A 3D surface rendering of object A from conical-beam lens data, taken in a pose where the plate is tilted with respect to the lens



(a)



(b)

Fig. 16 — Results of filtering the reconstructed surface shown in Fig. 15: (a) a low isosurface value, and (b) a high isosurface value

Figure 17 shows several different visualizations of the acoustic backscatter data from object D. In this image, the fin and one of the bolt patterns were not scanned and, hence, are not visible. The top-left panel is the traditional 2D range image, color coded red indicating near and blue indicating far. The range image is obtained by taking the first strong return (first return above a threshold) along each beam. If we take the first maximum return (above a threshold) along each beam, we get an image that is very similar. It can be seen that the cylindrical shape of the object is clear in the range image. However, the bolt pattern on the side of the cylinder, which has a finer scale, is not discernible. Also note that the strip of the cylinder that faces the lens is much better defined than are those parts of the cylinder that make an angle with the lens axis. The top-right panel is an *acoustic photograph* of the object. It is created from the intensity of the backscatter data. In this image, the bolts are quite visible because their concave corners form strong reflectors. Also note the image of the bottom in the cylinder and vice versa. The center panel is a surface-rendered 3D view of the gridded data. The noisy appearance of the cylinder is due to the fairly noisy range data, which becomes visually more acute when viewed in 3D. The lower-left panel is the 3D view shown at the center panel painted (or texture mapped) with the acoustic photo. In this image, the cylindrical surface appears less noisy, while at the same time the bolt heads are visible. Finally, the lower-right panel is the range data filtered so as to enhance small-scale features. Compared to the center panel, it can be seen that the cylindrical surface appears fuller, smoother and better defined, and the bolts stand out, albeit larger than they really are.

Figure 18 is similar to Fig. 17, except that the object pose during image acquisition was different. Note that in this image, both sets of bolts and nuts as well as the fin are visible.

In the case of objects resting on the bottom, the surface reconstruction for 3D rendering is complicated by interreflections between the object and the bottom. This object/bottom mirror effect produces false surfaces, e.g., the reflection of the bottom in the cylinder (Fig. 17) reaches the lens at a later time (than direct reflections) and is perceived by the imaging system as a true surface behind the cylindrical surface. The strengths of these interreflections are typically comparable to direct reflections from surface patches that face the lens, and are often stronger than reflections from surface patches that make an angle with the lens axis. Therefore, a surface detection algorithm could detect these false surfaces while missing the true surfaces. We have found that taking the first strong return (above a threshold) along each beam as the surface location, and discarding the remainder of the beam, can avoid this problem to some extent. However, it may happen that the return from a true surface patch with an unfavorable pose is not strong enough to be detected; hence, the range image will be a mixture of (mostly) true and (some) false surfaces, and thus appears noisy. The range images shown here are obtained in this manner.

Discussion

The acoustic images of suspended objects contain a great deal of detailed information. Small features of under 1 cm are often clearly visible. Object reconstruction and visualization is straightforward and require minimal image processing and filtering. Even when a surface patch makes an unfavorable angle with the lens, the backscatter energy is strong enough that it can be detected fairly accurately.

For targets on the bottom, the backscatter energy from the sandy bottom is comparable to the backscatter energy from the object, hence the low intensity returns from certain features of the object that would otherwise be clearly visible (if the object were suspended in the tank) become partially obscured. Moreover, the multiple reflections between the bottom and the object produce false surfaces. These two factors, the relatively weak returns from the object and the object/bottom mirror effect, result in missing surface patches or the appearance of ghost surface patches, thus making the task of reconstructing and interpreting the scene more challenging. To enhance small scale features

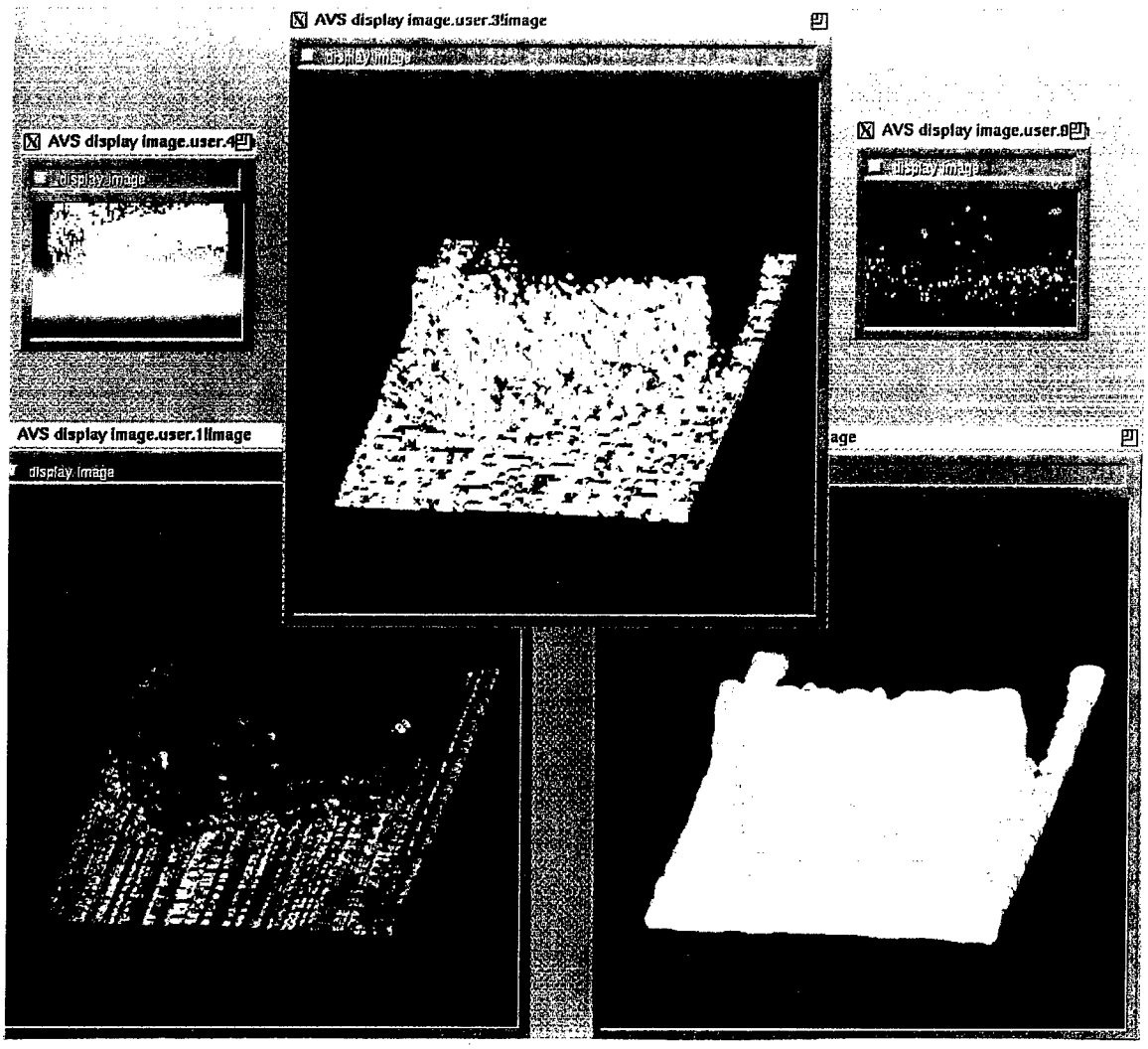


Fig. 17 — Acoustic images of object D lying on a sandy bottom taken by the conical-beam lens: (top left) range image (red is close, blue is far); (top right) acoustic "photograph," note the reflection of the bottom in the cylinder and vice versa; (center) surface rendering of the range data; (bottom left) a 3D view of the range data painted with the acoustic photo; (bottom right) a surface rendering of range data filtered so as to enhance smaller features. Note that the fin and one set of bolts are not visible in these images.

while removing the noise and smoothing surfaces, we have developed a set of volume-based filters that appear to perform well (see the Appendix). However, even better filters would be needed. Moreover, the issue of how best to present the data in this imaging scenario remains open.

CONCLUDING REMARKS

The images taken with the fan-beam lens under ideal conditions (i.e., with the target suspended in water) contain sufficient detail so that one can easily recognize familiar objects in them. In fact, a fan-beam imaging system for ship hull inspection (barnacle buildup detection) is being developed. However, the fan-beam images typically show less detail and contrast compared to those taken with the conical-beam lens and are more difficult to interpret. This is because the fan-beam has a wide angle in azimuth, which causes the returns to be averaged over a wider angle, thus reducing contrast.

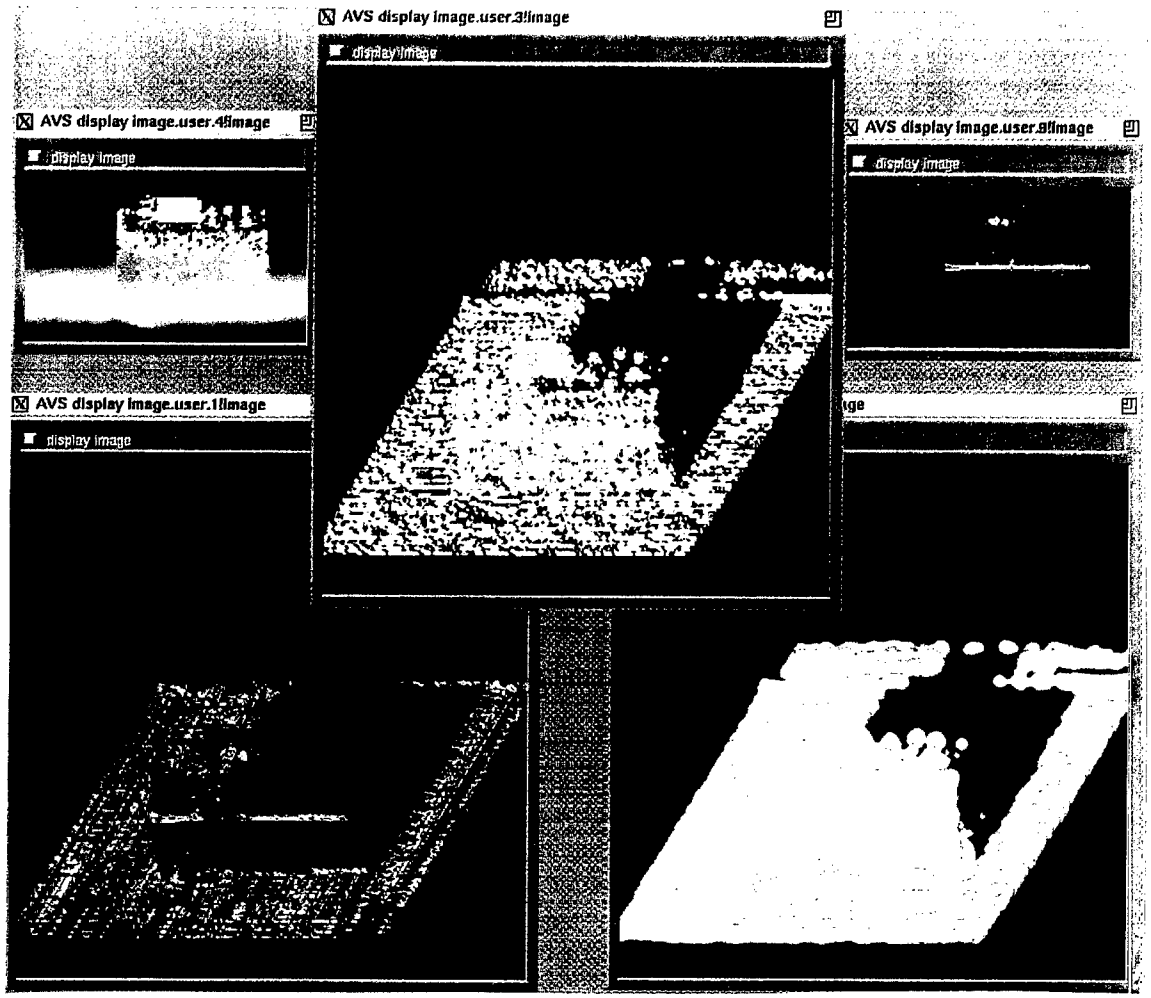


Fig. 18 — Similar to Fig. 17, except that the cylinder is in a different pose

It is possible to improve these images if the scene is imaged from many different directions. However, that introduces other problems such as sensor registration and longer times for image acquisition. The fan-beam image quality deteriorates significantly (to the point where they are not very informative) when the target is on the bottom of the tank because of the relatively strong bottom returns.

The images taken with the conical-beam lens, on the other hand, are highly detailed, vivid, and can be readily interpreted. Also, because the scene reconstruction and, hence, the images are 3D, they may be rotated in the computer and viewed from different poses. Conical-beam images contain considerable detail even when the target is lying on a sandy bottom, although they require some image processing to suppress noise and enhance small-scale features. Therefore, we believe that under field conditions, the conical-beam lens is the imaging system that will be the most useful.

The conical-beam imaging system prototypes used in this first phase of the work for acquiring the 3D images had a single transducer at the focal point of the lens. The images were acquired by mechanically scanning the scene. The resulting images have no cross-talk noise and no distortion due to off-axis aberrations. Thus, they represent the limit of what can be achieved with these imaging systems. To speed up image acquisition, two different approaches are currently under investigation. One approach is to populate the focal plane with many transducers (in conjunction with compound lens design to prevent off-axis aberrations), and the other approach is to use innovative scanning

strategies. Images from new prototypes are likely to be more noisy and should require a greater degree of image processing than we have had to use thus far. The scene reconstruction, object enhancement, and visualization aspects that need to be further investigated and improved are discussed in the following paragraphs.

Surface detection is a critical stage in surface rendering. We have thus far used gradient-based surface detection algorithms. There are a number of novel physics-based techniques for surface detection that may yield better results. These are based on surface particles that diffuse through the image field and/or interact with each other. These techniques are relatively slow and are typically used for interactive surface detection. Reference 15 provides examples.

The filtering algorithms that we have developed for surface smoothing, restoration, and small-scale feature enhancement are volume-based filters. There has been recent progress in surface-based filters, which may be more appropriate to use. Reference 16 describes one such filter.

Preliminary images from new prototypes typically show that features that produce strong backscatter also produce streaks and bleeding in neighboring pixels. A possible approach to dealing with these effects may be to use model-based filtering techniques.

The 3D acoustic images consist of sampled volumetric data. Volume rendering, rather than surface rendering, therefore, should be the most appropriate technique for visualizing the data. Recent progress in volume visualization techniques makes them more versatile and faster. These techniques, in particular VolVis [8], should be implemented and evaluated.

Finally, the high-frequency 2D array imaging system (the acoustic video), which is being developed, will output the range (i.e., first maximum return above a threshold) and the backscatter intensity. The range and intensity images so obtained contain complimentary information. The range image will show the large-scale structure of the object, while the intensity image will show strong reflectors, even though they may be small, such as corner reflectors. The question of how best to make use of these two sets of images is open. The possibilities include such techniques as displaying both images side by side, overlaying the two images (e.g., a 2D image shaded by range and colored by intensity), and wrapping the intensity image over a 3D image that has been reconstructed from range data. The range image obtained by the acoustic video may be quite noisy because of electrical and acoustical cross-talk, and because of mirror effects when objects are close to each other. It may be possible to use the intensity image to suppress range image noise while preserving small-scale features. However, we have to understand how interreflections affect the image before we can devise algorithms to exploit them. This may be done, for example, by extending the theory of shape from interreflections developed for optical images [17] to acoustic images, where interreflections play an important role.

ACKNOWLEDGMENTS

The imaging system prototypes were designed and tested by Ed Belcher (Applied Physics Laboratory, University of Washington), Don Folds (ARINC Research, Inc.), and Bruce Johnson (Naval Explosive Ordnance Disposal Technology Division). This project was sponsored by the Office of Naval Research.

REFERENCES

1. R.C. Gonzalez and P. Wintz, *Digital Image Processing*, 2nd edition (Addison-Wesley, Reading, MA, 1987).

2. J.C. Russ, *The Image Processing Handbook*, 2nd edition (CRC Press, Boca Raton, FL, 1994).

3. B.K.P. Horn, *Robot Vision* (MIT Press, Cambridge, MA, 1986).

4. B. Johnson, D. Scroggins, D. Folds, B. Kamgar-Parsi, and E. Belcher, "3-D Acoustic Imaging with a Thin Lens," Proc. IEEE Oceans '93 Conference, Victoria, BC, October 1993, vol. 3, pp. 444-449.

5. D.M. Scroggins and B. Johnson, "Test Report for High Resolution Acoustic Imagery Experiments," Technical Report TR-50-94-1, Naval Explosive Ordnance Disposal Technology Division, Indian Head, MD, July 1993.

6. L. Rosenblum and B. Kamgar-Parsi, "3D Reconstruction of Small Underwater Objects Using High-resolution Sonar Data," Proc. IEEE Symposium on Autonomous Underwater Vehicle Technology (AUV '92), Washington, DC, June 1992, pp. 228-235.

7. H. Fuchs, M. Levoy, and S.M. Pizer, "Interactive Visualization of 3D Medical Data," *COMPUTER*, pp. 46-51, August 1989.

8. R. Avila, T. He, L. Hong, A. Kaufman, H. Pfister, C. Silva, L. Sobierajski, and S. Wang, "VolVis: A Diversified Volume Visualization System," Proc. IEEE Visualization '94 Conference, Washington, DC, October 1994, pp. 31-38.

9. D. Lee, "Coping with Discontinuities in Computer Vision: Their Detection, Classification, and Measurement," *IEEE Trans. Pattern Analysis and Machine Intelligence* **12**(4), 321-344, 1990.

10. W.E. Lorensen and H.E. Cline, "Marching Cubes: a High Resolution 3D Surface Construction Algorithm," *Computer Graphics* **21**(4), 163-169, 1987.

11. B.T. Phong, "Illumination for Computer Generated Pictures," *Commun. ACM* **18**(6), 311-317, 1975.

12. A. Kaufman, *Volume Visualization* (IEEE CS Press, Los Alamitos, CA, 1991).

13. B. Kamgar-Parsi, L.J. Rosenblum, and E.O. Belcher, "Underwater Acoustic Imaging," *IEEE Computer Graphics and Applications* **12**(4), 11-13, 1992.

14. S.R. Marschner and R.J. Lobb, "An Evaluation of Reconstruction Filters for Volume Rendering," Proc. IEEE Visualization Conference, Washington, DC, Oct. 1994, pp. 100-107.

15. R. Szeliski, D. Tonnesen, and D. Terzopoulos, "Modeling Surfaces of Arbitrary Topology with Dynamic Particles," Proc. IEEE Conference on Computer Vision and Pattern Recognition, New York, NY, June 1993, pp. 82-87.

16. G. Taubin, "A Surface-based Smoothing Filter Without Shrinkage," Proc. International Conference on Computer Vision, Cambridge, MA, June 1995, pp. 486-491.

17. S. Nayar, K. Ikeuchi, and T. Kanade, "Shape from Interreflections," *Int. J. Computer Vision* **6**(3), 173-195, 1991.

Appendix

A FILTERING TECHNIQUE FOR NOISY SURFACES

This Appendix describes a sequence of operations that are designed to achieve two objectives. One is to reduce the noise level, i.e. smooth the surface, and the second is to enhance small features, i.e., enlarge them. We have found that a sequence of three image transforms (local averaging, high frequency emphasis filtering, and median filtering) performs satisfactorily. These are 3D extensions of 2D filters frequently used in digital image processing [A1].

A standard technique for smoothing is weighted local averaging, i.e., convolving the image with a Gaussian. The raw surface data Io is binary valued and one voxel thick.

When it is convolved with a 3D Gaussian mask G, the transformed image,

$$I_1(x,y,z)=G(x,y,z)*I_0(x,y,z),$$

becomes several voxels thick (depending on the mask size) and gray scale valued. In this work, we have used a size 3 mask. We have chosen the smallest size mask so as to perform a minimal amount of smoothing and to avoid losing small features.

High frequency emphasis (HFE) filtering, or the gradient image, in 2D intensity images is often used for edge enhancement. The gradient image in 3D,

$$I_2(x,y,z)=\left\|\vec{\nabla}I_1(x,y,z)\right\|,$$

however, has some features that are intuitively not obvious at first. In a gradient image, uniform intensity regions are transformed to zero and disappear. However, a binary-valued surface, say, a plane, forms a uniform region in two directions but presents a jump discontinuity in the third direction, which produces a nonvanishing gradient and preserves it in the image. Moreover, the surface becomes a pair of twin surfaces, one on either side of the original, displaced by one voxel. The HFE filter has the intended effect of enlarging small protrusions but makes holes appear smaller. In our work, we have used the $3 \times 3 \times 3$ gradient mask suggested in Ref. A2.

As we noted, the HFE filter transforms a surface into two surfaces. To remove this artifact, we apply a median filter,

$$I_3(x,y,z)=median\{I_2(x,y,z)\},$$

to the $3 \times 3 \times 3$ neighborhood of each voxel. The median filter has the desired effect of eliminating the twin surface artifact produced by the HFE filter and restoring the original surface to its correct position. The median filter also performs further smoothing of the noisy regions while preserving discontinuities.

The main effects of this sequence of image transforms can be observed by considering a simple example. Figure A1(a) is a surface rendering of a synthetic object composed of three planar surfaces with a hole and a bump. Figure A1(b) shows the result of the sequence of transforms. We can see that the one voxel thick surfaces have become thicker, the bump has become smoothed, although quite visible, and the hole is almost filled so that it appears like a concavity.

Figure A2(a) is the noisy synthetic object, where about 20% of surface voxels are noisy, i.e., displaced from their original position. The noise effects are rather severe in that neither the hole nor the bump are discernible. Figure 2(b) is the surface after filtering. Note that the filtered noisy surface is not very different from the filtered non-noisy surface, which indicates that these image transforms perform satisfactorily as far as removing noise and enhancing protrusions are concerned. The drawbacks are that holes become smaller and the surface appears thicker. Because the surface becomes thick and gray scale valued, we no longer have a single surface; rather, we have a family of isosurfaces. This gives us the opportunity to interactively render several isosurfaces to observe coarse to fine scale features. Low isosurface values correspond to smoother, less detailed surfaces.

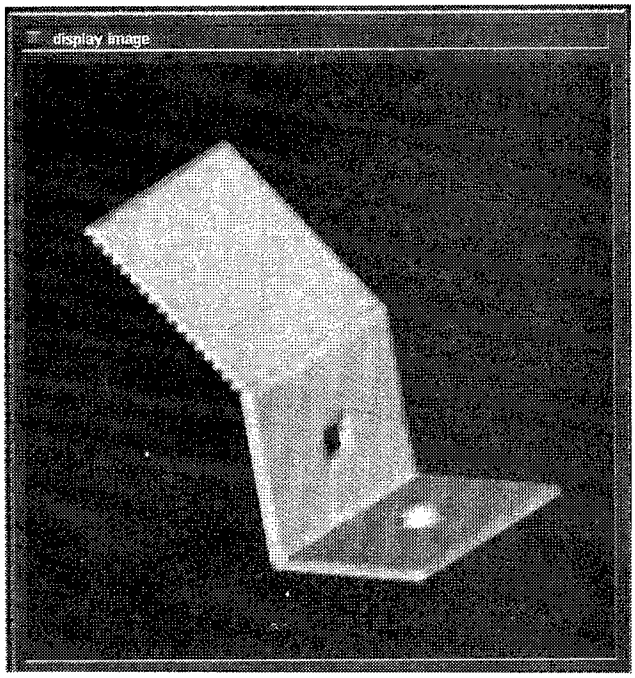
These filters need not be applied to the entire volume embedding the surface, but only to the voxels in the vicinity of the surface. Thus, the computational cost is $O(m)$, where m is the number of surface voxels, which can be significantly less than the number of volume voxels. They are also model free, unlike morphological filters, and, hence, applicable in a wide variety of situations. Reference A3 provides more details.

REFERENCES

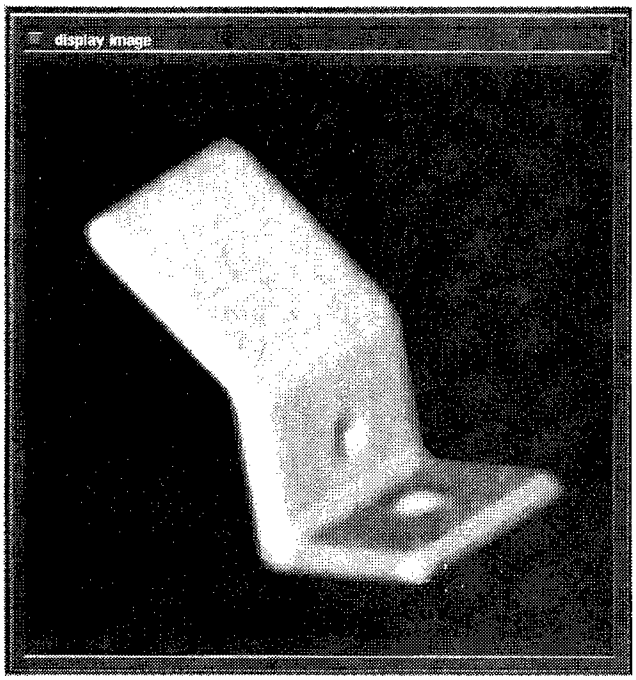
A1. A. Rosenfeld and A.C. Kak, *Digital Picture Processing* (Academic Press, NY, 1982).

A2. S.W. Zucker and R.A. Hummel, "A Three-dimensional Edge Operator," *IEEE Trans. Pattern Analysis and Machine Intelligence* **3**(3), 324-331, 1981.

A3. B. Kamgar-Parsi and B. Kamgar-Parsi, "Filtering Noisy Surfaces: Application to Underwater Acoustic Imaging," *IEEE Trans. Image Processing*, submitted 1995.

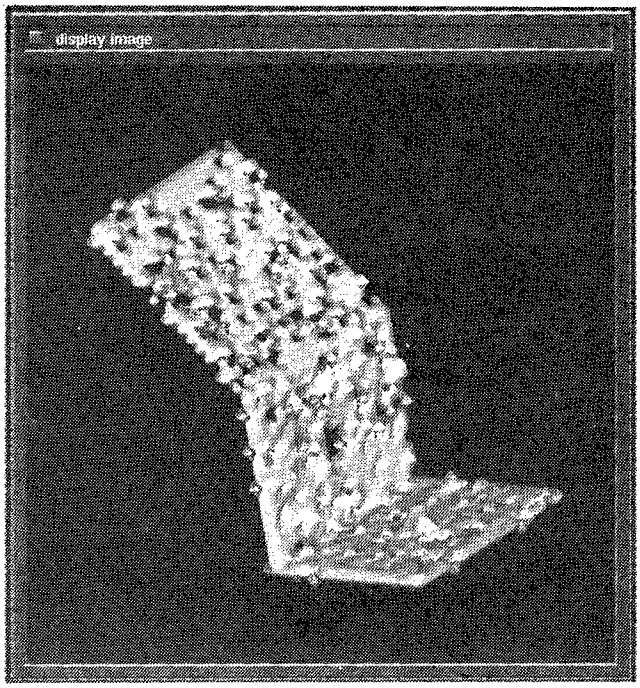


(a)

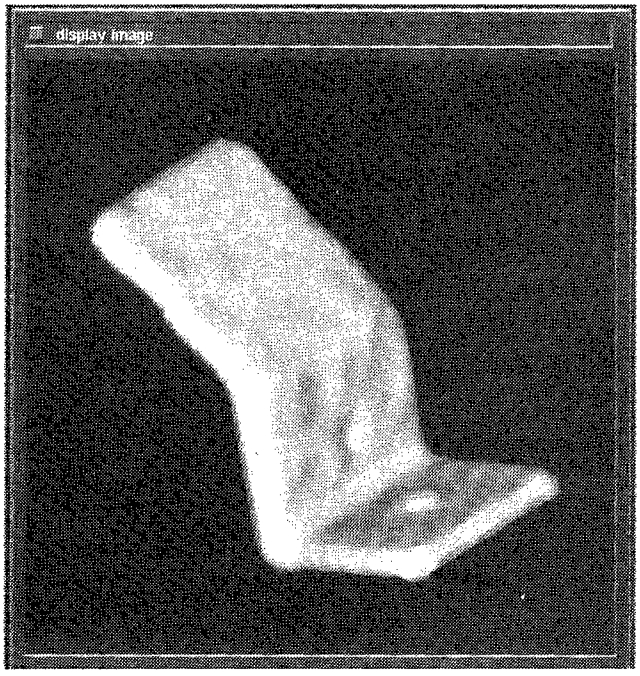


(b)

Fig. A1 — Isosurface rendered image of a noise-free synthetic surface before (a) and after (b) transforms



(a)



(b)

Fig. A2 — Isosurface rendered image of a noisy synthetic surface before (a) and after (b) transforms